

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 August 2002 (01.08.2002)

PCT

(10) International Publication Number
WO 02/059271 A2

(51) International Patent Classification⁷: C12N

(21) International Application Number: PCT/US02/02176

(22) International Filing Date: 25 January 2002 (25.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/263,757 25 January 2001 (25.01.2001) US
60/286,090 25 April 2001 (25.04.2001) US
60/292,517 23 May 2001 (23.05.2001) US

(71) Applicant (for all designated States except US): GENE LOGIC, INC. [US/US]; 708 Quince Orchard Road, Gaithersburg, MD 20878 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ORR, Michael, S. [US/US]; c/o GENE LOGIC, INC., 708 Quince Orchard Road, Gaithersburg, MD 20878 (US). NATION, Michele

[US/US]; c/o GENE LOGIC, INC., 708 Quince Orchard Road, Gaithersburg, MD 20878 (US). DIGGANS, James, C. [US/US]; c/o GENE LOGIC, INC., 708 Quince Orchard Road, Gaithersburg, MD 20878 (US). ZENG, Wen [CN/US]; c/o GENE LOGIC, INC., 708 Quince Orchard Road, Gaithersburg, MD 20878 (US).

(74) Agent: MORGAN, LEWIS & BOCKIUS LLP; TUSCAN, Michael S., WEIMAR, Elizabeth C. et al., 1111 Pennsylvania Avenue, N.W., Washington, DC 20004 (US).

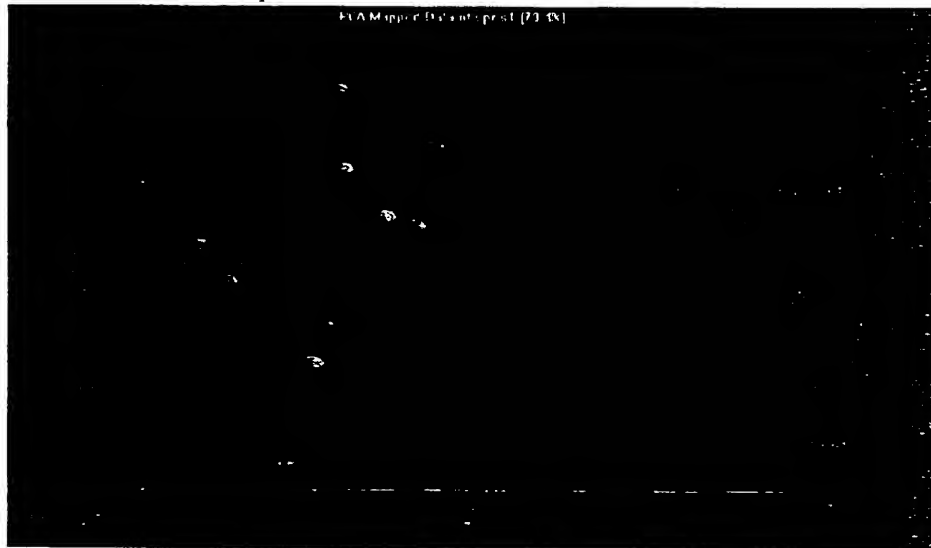
(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GI, GM, KB, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,

[Continued on next page]

(54) Title: GENE EXPRESSION PROFILES IN BREAST TISSUE

- 33 Tissue Samples



(57) Abstract: The present invention results from the examination of tissue from breast carcinomas to identify genes differentially expressed between tumor biopsies and normal tissue. The invention includes diagnostic and screening methods using these genes as well as solid supports comprising oligonucleotide arrays that are complementary to or hybridize to the differentially expressed genes.

WO 02/059271 A2

BEST AVAILABLE COPY

- (c) preparing second gene expression profile of the agent-exposed cell population; and
- (d) comparing the first and second gene expression profiles.

- 5 15. The method of claim 14, wherein the breast cancer is a infiltrating ductal carcinoma.
- 16. The method of claim 14, wherein the breast cancer is a microinvasive breast cancer.
- 17. A composition comprising at least two oligonucleotides, wherein each of the
10 oligonucleotides comprises a sequence that specifically hybridizes to a gene in Tables 1-5.
- 18. A composition according to claim 17, wherein the composition comprises at least 3 oligonucleotides.
- 15 19. A composition according to claim 17, wherein the composition comprises at least 5 oligonucleotides.
- 20. A composition according to claim 17, wherein the composition comprises at least 7 oligonucleotides.
- 20 21. A composition according to claim 17, wherein the composition comprises at least 10 oligonucleotides.
- 22. A composition according to any one of claims 17-21, wherein the oligonucleotides
25 are attached to a solid support.
- 23. A composition according to claim 22, wherein the solid support is selected from a group consisting of a membrane, a glass support, a filter, a tissue culture dish, a polymeric material, a bead and a silica support.
- 30 24. A solid support comprising at least two oligonucleotides, wherein each of the oligonucleotides comprises a sequence that specifically hybridizes to a gene in Tables 1-5.

25. A solid support according to claim 24, wherein the oligonucleotides are covalently attached to the solid support.
26. A solid support according to claim 24, wherein the oligonucleotides are non-
5 covalently attached to the solid support.
27. A solid support according to claim 24, wherein the support comprises at least about 10 different oligonucleotides in discrete locations per square centimeter.
- 10 28. A solid support according to claim 24, wherein the support comprises at least about 100 different oligonucleotides in discrete locations per square centimeter.
29. A solid support according to claim 24, wherein the support comprises at least about 1000 different oligonucleotides in discrete locations per square centimeter.
15
30. A solid support according to claim 24, wherein the support comprises at least about 10,000 different oligonucleotides in discrete locations per square centimeter.
31. A computer system comprising:
20 (a) a database containing information identifying the expression level in breast tissue of a set of genes comprising at least two genes in Tables 1-5; and
(b) a user interface to view the information.
32. A computer system of claim 31, wherein the database further comprises sequence
25 information for the genes.
33. A computer system of claim 31, wherein the database further comprises information identifying the expression level for the genes in normal breast tissue.
- 30 34. A computer system of claim 31, wherein the database further comprises information identifying the expression level for the genes in breast cancer tissue.
35. A computer system of claim 34, wherein the breast cancer tissue comprises infiltrating ductal carcinoma cells.

Reset

View

* Complete entries *

ID ABT11092 standard; cDNA; 2247 BP.
 XX
 AC ABT11092;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE Human breast cancer associated coding sequence SEQ ID NO: 1226.
 XX
 KW Human; breast specific gene; breast cancer; differential expression;
 KW cytostatic; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200259271-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 25-JAN-2002; 2002WO-US002176.
 XX
 PR 25-JAN-2001; 2001US-0263757P.
 PR 25-APR-2001; 2001US-0286090P.
 PR 23-MAY-2001; 2001US-0292517P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Orr MS, Nation M, Diggans JC, Zeng W;
 XX
 DR WPI; 2002-674803/72.
 XX
 PT Diagnosing breast cancer in a patient comprises detecting the level of
 PT gene expression in cell or tissue samples, where a differential gene
 PT expression is indicative of breast cancer.
 XX
 PS Claim 1; SEQ ID NO 1226; 260pp + Sequence Listing; English.
 XX
 CC The present invention relates to methods of diagnosing breast cancer in a
 CC patient, which comprise detecting the level of expression in a tissue
 CC sample of two or more genes selected from those shown in ABT09867-
 CC ABT11112, where a differential expression of the genes indicates breast
 CC cancer. The methods are useful in diagnosing, treating, detecting the
 CC progression, and in monitoring treatment of breast cancer in patients.
 CC The methods are also useful as a screening tool for agents that modulate
 CC the onset or progression of breast cancer. The breast cancer genes may be
 CC used as diagnostic markers for the prediction or identification of the
 CC malignant state of breast tissue, for confirming the type and progression
 CC of cancer, and for drug screening and assays. The present sequence is a
 CC coding sequence of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub.published_pct_sequences
 XX
 SQ Sequence 2247 BP; 393 A; 715 C; 675 G; 464 T; 0 U; 0 Other;
 ccggggcgga tggctccggc cgctggctc cgcagcgcg cgcgcgcgc cctcctgccc 60
 ccgatgctgc tgctgctgct ccagccgccc ccgctgctgg cccgggctct gccgccggac 120
 gtccaccacc tccatgccga gaggaggggg ccacagccct ggcattgcagc cctgcccagt 180
 agcccggcac ctgcccctgc cagcaggaa gcccccggc ctgccagcag cctcaggcct 240
 ccccgctgtg gcgtgcccga cccatctgat gggctgagt cccgcaaccg acagaagagg 300

QUEST AVAILABLE COPY

ttcgtgcttt	ctggcgggcg	ctgggagaag	acggacctca	cctacaggat	ccttcgggttc	360
ccatggcagt	tgggtgcagga	gcaggtgcgg	cagacgatgg	cagaggccct	aaaggatatgg	420
agcgatgtga	cgccactcac	ctttactgag	gtgcacgagg	gccgtgctga	catcatgatc	480
gacttcgcca	ggtactggca	tggggacgac	ctgccgtttg	atgggcctgg	gggcatcctg	540
gccccatgct	tcttccccaa	gactcaccga	gaaggggatg	tccacttcga	ctatgatgag	600
acctggacta	tcggggatga	ccagggcaca	gacctgctgc	aggtggcagc	ccatgaattt	660
ggccacgtgc	tggggctgca	gcacacaaca	gcagccaagg	ccctgatgtc	cgccttctac	720
acctttcgct	acccactgag	tctcagccca	gatgactgca	ggggcggttc	acacctatat	780
ggccagccct	ggcccaactgt	cacctccagg	acccagccc	tgggccecca	ggctgggata	840
gacaccaatg	agattgcacc	gctggagcca	gacgccccgc	cagatgcctg	tgaggcctcc	900
tttgacgcgg	tctccaccat	ccgaggcgag	ctctttttct	tcaaagcggg	ctttgtgtgg	960
cgcctccgtg	ggggccagct	gcagcccggc	taccagcat	tggcctctcg	ccactggcag	1020
ggactgcccc	gccctgtgga	cgctgccttc	gaggatgccc	agggccacat	ttggttcttc	1080
caaggtgctc	agtactgggt	gtacgacggt	gaaaagccag	tcctggggcc	cgcacccctc	1140
accgagctgg	gcctggtgag	gttcccggtc	catgctgcct	tggctctggg	tcccgagaag	1200
aacaagatct	acttcttccg	aggcagggac	tactggcggt	tccaccccag	caccgcgcgt	1260
gtagacagtc	ccgtgccccg	cagggccact	gactggagag	gggtgccctc	tgagatcgac	1320
gctgccttcc	aggatgctga	tggctatgcc	tacttctctg	gcggccgcct	ctactggaag	1380
tttgaccctg	tgaagggtgaa	ggctctggaa	ggcttcccc	gtctcggtgg	tcctgacttc	1440
tttggtctgt	ccgagcctgc	caacactttc	ctctgacct	ggcttggatg	ccctcagggg	1500
tgctgacccc	tgccaggcca	cgaatatcag	gctagagacc	catggccatc	tttgtggctg	1560
tgggcaccag	gcatgggact	gagcccatgt	ctcctgcagg	gggatggggt	ggggtacaac	1620
caccatgaca	actgccggga	gggccacgca	ggctgtggtc	acctgcagc	gactgtctca	1680
gactgggcag	ggaggctttg	gcatgactta	agaggaagg	cagtcttggg	acccgctatg	1740
caggctcctg	caaacctggc	tgcctgtct	catccctgtc	cctcagggta	gcaccatggc	1800
aggactgggg	gaactggagt	gtccttgctg	tatccctgtt	gtgaggttcc	ttccaggggc	1860
tggcactgaa	gcaaggggtg	tggggcccca	tggccttcag	ccctggctga	gcaactgggc	1920
tgtagggcag	ggccacttcc	tgaggtcagg	tcttggtagg	tgcctgcac	tgtctgcctt	1980
ctggctgaca	atcctggaaa	tctgttctcc	agaatccagg	ccaaaaagtt	cacagtcaaa	2040
tggggagggg	tattcttcat	gcaggagacc	ccaggccctg	gaggctgcaa	catacctcaa	2100
tctgtccca	ggccgatcc	tcctgaagcc	cttttcgcag	cactgctatc	ctccaaagcc	2160
attgtaaatt	tgtgtacagt	gtgtataaac	cttcttcttc	tttttttttt	ttaaactgag	2220
gattgtcatt	aaacacagtt	gttttct				2247

//

BEST AVAILABLE COPY